4 PERFORMANCE REQUIREMENTS

SFA's modernized systems will likely be implemented through many different hardware and software combinations, involving a number of different contractors. Each mechanism and/or contractor will have an individual set of performance requirements. The intent of this Section is to provide SFA with a consistent set of performance requirements that apply to mechanisms and contractors. The same requirements will apply to anyone who develops a part of EASI/ED.

The performance requirements were developed by examining current Title IV systems procedures, leveraging industry best practices, and identifying transactions that are performance sensitive. Identification of these transactions indicates where critical performance issues may occur, and consequently, slow the financial aid delivery process. At the same time, examining the current Title IV systems establishes a baseline of current practices upon which to base future recommendations. Industry best practices, where applicable, are combined with current Title IV best practices to bring validation to the conclusions drawn. This Section comprises the following subsections:

- Subsection 4.1 Introduction to performance requirements
- Subsection 4.2 Assumptions
- Subsection 4.3 High-Volume and Performance Sensitive Transaction Classes
- Subsection 4.4 Performance Requirements
- Subsection 4.5 High-level Performance Issues for Reused Title IV Systems

4.1 Introduction to Performance Requirements

This subsection describes the purpose and meaning of performance requirements. It also defines two important considerations, backup and recovery, that must be combined with quantitative performance measures to ensure optimum performance.

According to Faulkner Information Services, the objectives of performance requirements are to:

- optimize the use of systems and resources;
- effectively supervise data processing activities;
- identify existing and/or potential data management and transaction processing issues;
- identify opportunities to fine-tune system operations;
- plan and implement long-term system upgrades; and
- ensure data security, safety, and availability.

Performance measures used to capture system performance include:

- **Availability** – this is a measure of the system being ready for user activity. An available system provides services immediately to users, whereas a system that suffers from low availability typically forces users to wait.

- **Response Time** – this measures the system reaction time to a given input. In other words, this is the time interval between pressing the send key of a terminal and the display of the system's reply.

- **Transaction Volume** – this quantifies the number of transactions completed within a specified period. One year is the period used in this analysis. While important, volume alone is not a valid indicator of computer performance. Performance also considers the number of user requests handled, as well as the degree of responsiveness to those requests. Regular patterns in transaction volume help predict peak resource usage times, and shifts in transaction volumes indicate new trends not previously evident. For example, it is clear that disbursement processing occurs heavily in January, indicating peak resource usage at that time.

- **Reliability** - this is a measure of a system's ability to continue operations despite the failure of some critical element, and credibility of the information provided. Typically, reliability involves redundancy, which can increase component and media costs.

Two other considerations that need to be addressed are backup and recovery. Although generally not measured in quantitative terms, if data is lost and cannot be brought back, the performance of the system will suffer accordingly. These two considerations can be defined as the following:

- **Backup** – this is the process of copying all the files on a system to another storage media so that they will be preserved in case of equipment failure or other catastrophe.

- **Recovery** – this is the method of resuming operations following a hardware or software failure and subsequent restoration of service.

The above measures and considerations are critical issues in defining EASI/ED performance. These elements are necessary to determine and evaluate performance levels for servers, clients, applications, and networks, as well as to set baselines and diagnose problems. System downtime, slow response times, bottlenecks during peak usage periods, and limited bandwidth can frustrate users as well as reduce the overall quality of service. Performance requirements are used to set standards that aim to minimize these problems.

4.2 Assumptions

This subsection documents assumptions about the EASI/ED environment and users that influenced the performance requirements analysis.

1. EASI/ED partners and customers access EASI/ED functionality through the Internet and/or direct connections. Most users of the EASI/ED system (e.g. participants, lenders, guarantors, and state grant agencies) access the system remotely. Schools and Aid Organizations have remote access, as well as direct connection to EASI/ED. ED personnel primarily accesses the system locally, however, they may also have remote access accounts.

2. The transactions that need to be analyzed for performance sensitivity are the data flows defined in the *Project EASI/ED ASDD, version 2*.0.

3. The hours of 8:00 AM to 8:00 PM, Eastern Standard Time (EST) will represent the busiest time of the day for user access to EASI/ED.

4. Analog modems will be the primary means of access to the Internet for the next 3-5 years.

5. FFELP loan information will continue to process outside of EASI/ED, with reporting information fed into EASI/ED later. Direct loan partners and customers will be the predominant users of EASI/ED.

6. Performance requirements will not be any less stringent than those in place for existing Title IV systems.

7. Transactions in EASI/ED are processed both online and through batch processing. Online transactions are completed and updated interactively while accessing the system, whether that access is remote or direct connection. Batch transactions are queued over a period, then processed at a specific time.

8. Transaction volumes will continue to grow as noted in the comments column of Appendix F of the *Project EASI/ED ASDD, version 2.0*, unless otherwise noted.

9. Scheduled downtime does not impact availability measurements.

10. System availability is based on an average capacity of 30 days per month and 24 hours per day, or 720 hours per month and 8640 hours per year.

4.3 High-Volume and Performance Sensitive Transaction Classes

The intent of this analysis is to identify transactions that are significantly performance sensitive by nature. A **transaction** refers to the completion of a specific unit of work. For example, a request for income verification is a transaction that EASI/ED sends to the Internal Revenue Service (IRS). Determining these resource-intensive transactions helps identify potential trouble spots and bottlenecks in the delivery of financial aid. Performance requirements should be developed that will address these transactions. As a result, only those transactions deemed performance sensitive by nature are listed.

Based upon data and information gathered from industry best practices, EASI/ED Joint Information Gathering (JIG) sessions, and EASI/ED requirements, performance sensitive transaction classes associated with implementing Project EASI/ED have been categorized into four areas. These areas are discussed in the following subsections:

- Subsection 4.3.1 Online updates with high-volume activity
- Subsection 4.3.2 Transactions with long and complex access paths
- Subsection 4.3.3 Transactions that access entity types with a combined high level of activity
- Subsection 4.3.4 Groups of transactions that execute together and impact each other

The following subsection details the approach taken to identify each type of performance sensitive transaction. Through the course of the analysis, it was determined that the transactions identified in subsections 4.3.3 and 4.3.4 were not sufficiently selective in identifying performance sensitive transactions. Therefore, transactions identified exclusively using either of those two

methods were eliminated. Appendix F captures all remaining transactions after the filtering process is applied.

4.3.1 Online Updates with High-Volume Activity

High volume online transactions often produce bottlenecks during transaction processing and can cause users to experience slower response times. Transactions that use some type of online processing (i.e., online query, online with batch update, or online real-time update) are identified and categorized into the following volumes:

1. Low – anticipated yearly transaction volume less that 40,000.

2. Medium – anticipated yearly transaction volume greater than or equal to 40,000 and less than 1 million.

3. High - anticipated yearly transaction volume greater than or equal to 1 million.

The 17 transactions with volumes of one million or greater account for 99 percent of the total high volume online transactions. The periods of the first and fourth quarters are especially volume sensitive, as many of the transactions occur during this time.

4.3.2 Transactions with Long, Complex Access Paths

A transaction may access multiple data entity types during the course of its processing. As transactions access multiple entity types and access these entity types more than once, the path taken by the transaction can become extremely long and complex, causing strains on system performance. The following steps were executed to identify the transactions that fit into this category:

1. Analyze all data flows identified in the *ASDD,* and consider each data flow one transaction.

2. Identify the entity types accessed by each transaction.

3. Total the number of effected entity types by transaction.

The most complex transactions, the top 26 of the 389 transactions listed, affected over 7 times as many entity types as the remaining transactions. Appendix F includes these transactions.

4.3.3 Transactions that Access Entity Types with Combined High Level of Activity when Accounting for All Transactions Against the Entity Type

Across EASI/ED subsystems, many entity types will be involved in multiple transactions in order to eliminate duplicate data entry and maximize database performance. An entity type is anything about which the system must store data (e.g. a person, place, thing, event, or concept). An example for EASI/ED is “School”. This circumstance drives up the “demand” on these entity types, which affects system performance. The purpose of this subsection is to identify the entities accessed most, then list the transactions that access these entity types. The following steps were conducted to identify the transactions that access the “high traffic” entity types:

1. Determine the total number of transactions that access or “hit” each entity;

2. Identify the ten most frequently accessed entity types; and

3. List each transaction that accesses at least one of the top ten entity types.

Two thousand three hundred and forty seven hits occur across all 106 entity types. Approximately 50% of the activity includes just ten entity types:

- AID
- AID_ACCOUNT_TRANSACTION
- AID_ORGANIZATION
- AID_ORGANIZATION_POC
- AID_PROGRAM
- PARTICIPANT
- PARTICIPANT_ADDRESS
- PARTICIPANT_NAME
- SCHOOL
- SOCIAL_SECURITY_NUMBER

Increased traffic across these entity types can slow response times due to bottlenecks. This analysis yielded 281 transactions that hit each of the ten highest activity entity types. Given that this number represents 72% of all transactions, it is concluded that this particular category does not discriminate enough in identifying individual performance sensitive transactions.

4.3.4 Groups of Transactions that Will Execute Together and Impact Each Other

When processed individually, most transactions are not overly resource intensive. However, financial aid transactions often process at the same time during the academic year as the financial aid process moves forward.

Listed in Table 4 - 1 below, are seven transaction categories and their anticipated peak EASI/ED processing periods. For each column (i.e., processing period that has multiple transaction category boxes checked) the implication is that the transactions occurring in that time period are subject to executing together. The columns begin with the month of July.

Transaction Category	J	A	S	O	N	D	J	F	M	A	M	J
Loan Origination	X	X	X									
Loan Disbursement		X	X	X	X		X	X				
Enrollment Tracking			X					X				
Participant Authorization			X	X	X		X	X				
Interest and Special Allowance	X			X						X		
Default Rates				X			X	X	X			
Financial Institution Reporting			X			X			X			X

Table 4 - 1: Groups of Transactions that Execute Together and Impact Each Other

Analogous to a highway during rush hour, a system with a great deal of transaction traffic tends to suffer in performance. The large numbers of transactions alone affect the performance of the system, but when combined with other transactions occurring at the same time, it serves to cause immense pressure on the system. As evidenced in the above table, in the course of the delivery of financial aid, there are several periods, and respective transactions, that process on or near the same time. These transactions have the potential to be performance sensitive.

Under this approach, 60% of the 389 transactions were identified as falling into this category. This implies that there are some periods of very heavy system activity, such as January in the first quarter, and September and October in the third and fourth quarters. However, it was determined that this method does not discriminate enough to determine individual performance sensitive transactions.

4.4 Project EASI/ED Performance Requirements

This subsection includes performance requirements for Project EASI/ED grouped into eight categories:

- Subsection 4.4.1 Concurrent User Requirements
- Subsection 4.4.2 Response Time Requirements
- Subsection 4.4.3 System Requirements
- Subsection 4.4.4 Data Management Requirements
- Subsection 4.4.5 Backup Requirements
- Subsection 4.4.6 Recovery Requirements
- Subsection 4.4.7 Availability Requirements
- Subsection 4.4.8 Reliability Requirements

4.4.1 Concurrent User Requirements

An estimate of the projected number of concurrent users that require support for the user types defined in section 3, Security Requirements, is based on current systems data and Project EASI/ED functional requirements.

The postsecondary community can be segmented into two primary designations: customers and partners. Customers include over fourteen million graduate, undergraduate and other postsecondary students, and their families. Partners include over 6,200 postsecondary schools and institutions, 36 guaranty agencies and state grant agencies, and 6,500 lenders. Finally, ED personnel also require access and use of the system. Given the tremendous numbers of potential users, it is important to estimate the number of users who will be using the system simultaneously.

Although a significant number of users will access the system through local means, most users will access the system primarily through remote access computers (i.e. through the Internet). Based upon the information gathered in the high volume transaction subsection, as well as other ED provided information, Table 4 – 2 represents estimates regarding the volume of users that will be using the system concurrently.

It is necessary to incorporate some assumptions when building these figures, due to the lack of solid historical data. Although grounded in logic, these assumptions must be kept in mind when regarding these concurrent usage requirements. It is important to note that the estimate of the concurrent users could change radically if changes are made in the assumptions. This being the case, these estimates should be viewed as guidelines to build around, rather than specific numbers that must be accommodated.

User Type	Application	Origination/ Disbursement	Repayment	Financial Services	Prgm. Mgmt. Oversight
Customers	8,000[1]	1,148[2]	4,593[4]	Not Applicable	Not Applicable
Schools	Minimal	1,808[3]	Not Applicable	Not Applicable	477[7]
Financial Institutions	Minimal	Not Applicable	Not Applicable	Not Applicable	296[8]
ED	Not Applicable	Minimal	35[5]	10[6]	24[9]
Total	8,000	2,956	4,628	10	597
Population Comments	Percent of participant population that files the FAFSA over Internet; no schools, FI, or ED substantially involved.	Only participants receiving direct loans; only direct loan schools; no financial institutions; ED not substantially involved.	Participants repaying direct loans; schools and FI not involved (direct loans only). Debt collection ED population eligible.	Participants, schools, and FI not involved; AFMS ED staff population eligible.	Participants not involved. Entire Schools, FI, and APOS and GLOS ED population eligible.

Table 4 – 2: Concurrent User Estimates

The concurrent usage numbers indicated in the above table represent estimates for Fiscal Year 2001. The assumptions and calculations used to derive these numbers can be found at the back of Appendix F. From a high level standpoint, the peak concurrent usage period is during the application process, due to the significant access needed by prospective students completing the FAFSA on the web. The disbursement and origination period is also somewhat busy, again influenced significantly by heavy participant access, as well as very heavy school access. Although repayment is not particularly time sensitive, given the large number of participants expected to be in repayment at any one time, it is expected to have a large number of concurrent users. Since participants are not involved in financial services, program management, and oversight, these sections have relatively small concurrent usage requirements.

4.4.2 Response Time Requirements

Response time requirements are based on the functional requirements documented in the *Project EASI/ED BARD, version 2.0*, and incorporate industry best practices. The response time requirements are built with the intention of satisfying the processing needs of the high volume online transactions, identified in Appendix F. These requirements consider access types such as local, remote, and mobile and recommend reasonable maximum response times in each case. They also address the issue of accessibility and include a review of Title IV response requirements.

Response time is the elapsed time between the entry of an inquiry or demand on a computer system and the returned response from the system. An example of this is the time between hitting the *Enter* button upon entry of a query, and having the query's answer displayed on the user's terminal. There is also the concept of *perceived* response time, which is the time a user senses as the beginning of input and the end of the response. Any delay or waiting that increases real or perceived response time beyond the target baseline is referred to as latency. As a component of performance measures, response time is critical because it directly measures end user satisfaction. However, it is also one of the most difficult components of performance to measure and forecast due to the fact that many applications require information from several different programs in order to complete a single request. This makes it difficult to pinpoint the specific problem area even when an unsatisfactory transaction response time occurs. The problem could be the network, server, client, application, location of the information, or any combination of these factors. The best approach to forecasting is to identify acceptable targets and maximums based on the average response time in a variety of environments and configurations.

Response time can also be affected by the method of accessing information. For this analysis, the type of access that EASI/ED users have is through either a Local Area Network (LAN) or Wide Area Network (WAN), or some sort of remote transmission, commonly through the Internet. These access types are defined as follows:

- **Dedicated Network - LAN/WAN Access -** a LAN is a network of interconnected workstations sharing the resources of a single processor or server within a relatively small geographic area. Typically, this might be within the area of an office building. When a user has direct access to the LAN, he or she connects to the network through a cable that typically transmits data at a rate of 10 or 100 Millions of bits per second (Mbps). Mbps is a measure of bandwidth that translates into the total information flow over a period of time. A WAN is a generic term for any network that covers a large geographic area (50 miles, 80 km, or greater) and includes packet-switching, public data, and value-added networks.

- **Remote/Mobile Access** - remote access offers the ability to gain access to a computer or a network from a remote location. Typically, these users will access the Internet through an Internet Service Provider (ISP) which they connect to using a modem. Although remote access is also possible using other transmission methods, a dial-up Plain Old Telephone Service (POTS) transmission using a computer modem is the most common method of gaining access to the Internet and other remote sites. A POTS connection currently only transmits data through a 14.4, 28.8, or 56 thousands of bits per second (Kbps) modem.

Current Title IV System Response Times

The response time for current Title IV systems were examined to help project the appropriate response times for EASI/ED. When examining current Title IV Systems and their PWSs, the maximum response time was not to exceed 15 seconds (Table 4 - 3). When accessing a system through a dedicated LAN connection, the requirements generally dictate that it be 75 percent faster than through remote access. Again, referring back to the current Title IV Systems PWSs, the maximum direct LAN connection access time is not to exceed four seconds (Table 4 – 3).

Title IV System	Remote Access Max Response Time (seconds)	LAN Maximum Response Time (seconds)	Comments
LSS	15	4	Response times stated in system contract.
CPS	Not Available	4	Dedicated LAN response time stated in system contract. Remote time not listed.
NSLDS	15	4	Response times stated in system contract.
TIV WAN	12	Not Available	"Access time" of 12 seconds stated in contract. FIPS PUB 144 referenced.
FFELP	Not Available	4	Dedicated LAN response time stated in system contract. Remote time not listed.
LOS	15	4	Response times stated in system contract.
MDE	Not Available	Not Available	No response times specified in the contract.
CDS	Not Available	Not Available	No response time specified in the contract.
RFMS	Same as TIVWAN	Same as TIVWAN	Stated in contract that RFMS shall comply with telecommunications standards covering the TIVWAN.
PEPS	Not Available	Not Available	No response times specified in the contract.
CBS	Not Available	Not Available	No response times specified in the contract.

Table 4 – 3: Current Title IV Systems Contractual Response Times

In summary, the average maximum response times were 13.8 seconds through remote access and 4 seconds through the LAN.

Benchmark Response Times

Response time data submitted to The Standard Performance Evaluation Corporation (SPEC) was reviewed for applicability to SFA's requirements. SPEC is a non-profit corporation with the mission to "establish, maintain and endorse a standardized set of relevant benchmarks that can be applied to the newest generation of high-performance computers". While no one benchmark can fully characterize overall system performance, the results of a variety of realistic benchmarks can give valuable insight into expected real performance. SPEC conducted a study of response times for system file servers (SFS) and web servers. When analyzing this data, the systems that report the highest and lowest times were not considered, in order to prevent any possible skewing of averages due to extreme values. The response times for SFSs was found to average five seconds. The maximum SFS response time referenced is 36.3 seconds. The average Web Server response time was 7.43 seconds. The maximum Web server response time considered was 28.7 seconds.

Based upon these results, the following numbers are recommended as the maximum response times when the system is normally loaded, not during extreme peak periods:

- When the system is normally loaded, it shall ensure that remote users receive a maximum response time of 7.5 seconds 95% of the time, and never exceed a response time of 29 seconds; and

- When the system is normally loaded, it shall ensure that local users receive a maximum response time of 4 seconds 95% of the time, and never exceed a response time of 36 seconds.

4.4.3 System Performance Requirements

The Financial Aid process is cyclical in nature. Applications, as entered via the FAFSA, will be followed by origination, which is followed by disbursement, and so on. Each general process occurs at similar times each year. For example, as seen in Figure 4 - 1, disbursements for direct loans have a large peak in the third and fourth quarter of the year, as the Fall semester begins, and also have a significant peak in January, due to the start of the Spring semester. This cyclical distribution is generally the case for each type of transaction involved in the delivery of financial aid.

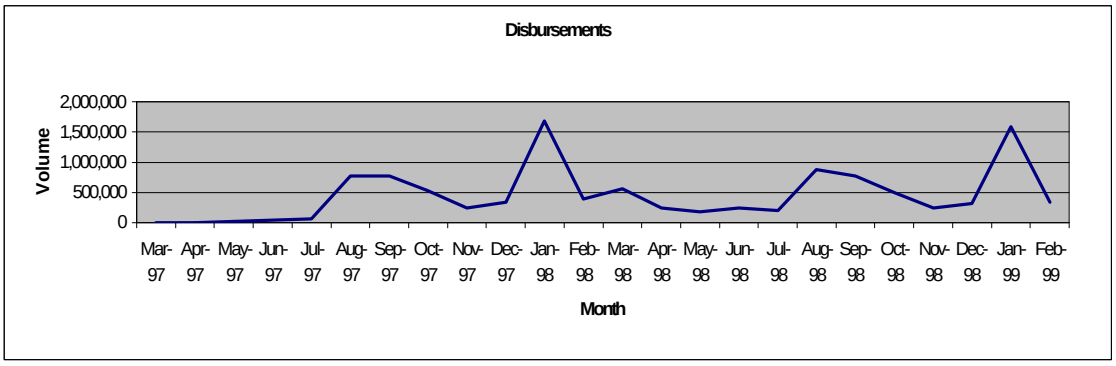

Figure 4 -1: Direct Loan Disbursement

As part of the performance requirements for EASI/ED, it is important to determine the volume of these transactions, and the amount of time necessary to complete these transactions. Turnaround time is defined as the length of time it takes for these transactions to be completed. For many of these transactions, ED and other members of the postsecondary community have established a turnaround, or transaction completion, time. For example, a FAFSA, once received, needs to be processed and completed within 8 days. The volume of transactions that the system must accommodate also has a distinct impact on the performance and service level of the system. Volume is defined as the total number of occurrences of that transaction within one calendar year. As seen in the direct loan disbursement graph, most transactions will tend to occur predominantly within specific periods during the year.

The turnaround times and volumes for performance sensitive transactions are listed in Appendix F. Any peak periods for individual transactions are also noted.

4.4.4 Data Management Requirements

Data Management requirements are built on assumptions in the *Project EASI/ED Transition Strategy* (September 25, 1998) and the *Project EASI/ED LDMD* and address replication and best practices regarding its use. Data may be *replicated*—periodically copied from a central database to other databases—in order to minimize the risk of central database failure, and to make the data available to as many users who require it in a reasonable amount of time.

Data replication occurs when data is copied from a central database to multiple target databases. From a performance perspective, the purpose of replication is to bring the data contained within the initial database to a location closer to the prospective user. Please see Figure 4 - 2. Similar to transactions that execute together, users who access a database at the same time create heavy user traffic, and the performance of the system suffers accordingly. Replicating the data in EASI/ED into additional databases (three, in the example) allows more users to access the same information at the same time, without suffering the performance liabilities inherent in access to a single database.

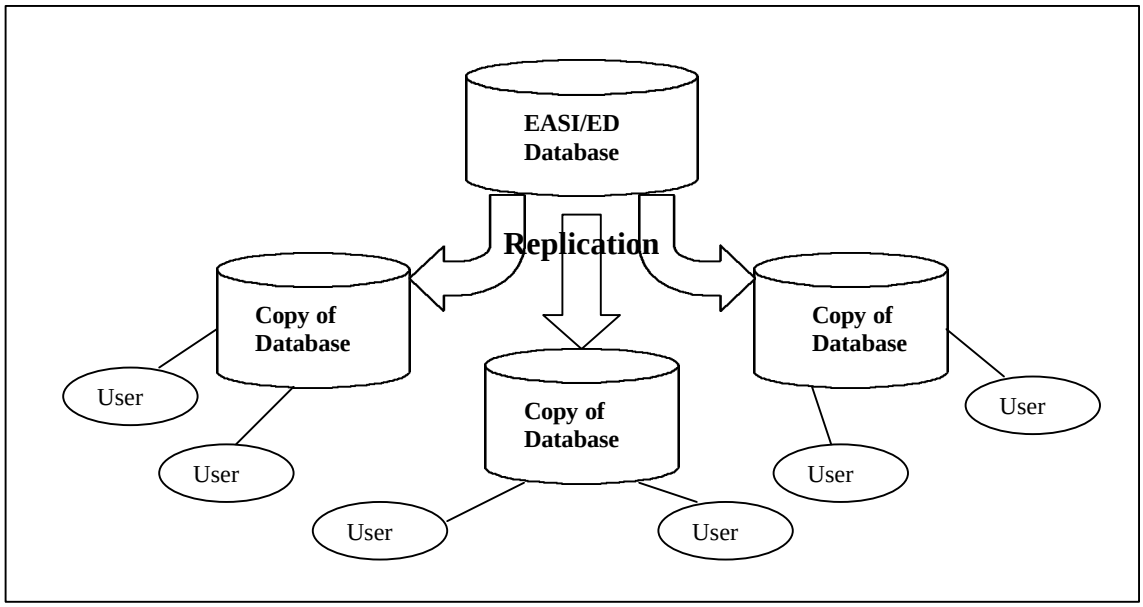

Figure 4 -2: Replication

Data conflicts can occur during replication. A conflict occurs when some change, be it a deletion, insertion, or update, is made at a local database and is posted to the master database at the same time a change was made to the record at the master database. As a result, centralized business rules, or reconciliation procedures, are required in order to handle any discrepancies occurring as a result of the replication process.

It has also been found that as the number of target sites for replication increases, so does the likelihood for data inconsistency. Best practices indicate that the target sites should be limited to fewer than ten, with the ideal levels being less than five. Any more than ten will simply cause too much opportunity for the creation of data inconsistency

The following data replication requirements are recommended:

- Centralized business rules should be established and maintained to handle conflicts in replication; and

- Target sites for replication should be kept to fewer than ten sites, with the ideal levels being under five sites.

4.4.5 Backup Requirements

Data loss, however minor, can seriously damage an organization's ability to survive if ineffective or insufficient backup procedures were in place before the disaster. In addition to sabotage, vandalism, equipment failures, and intermittent power outages, environmental disasters - floods, fires, tornadoes, and hurricanes - are an ever-present threat to data and systems security. In order to ensure that data is never completely lost, and to facilitate rapid recovery of operations when data is lost, backup procedures should be established.

The method and frequency of backups will vary, but best practices indicate that all of the following are used in various combinations:

- **Full Backup** – This occurs when all the data contained within the system is copied. This type of backup is often done on a weekly basis, with some organizations finding it necessary to conduct it on a daily basis.

- **Partial Backup** – This occurs when only critical data held on the system is copied. As this backup only captures mission critical information, it is done more often than a full backup, since it takes less time and space. A full backup is then conducted on a routine basis to capture all the data.

- **Incremental Backup** – This occurs when only data that has been affected since the last backup is copied. Similar to the partial backup, this method does not collect all data present on the system, only transaction files and data that has changed. As with the partial backup, this method is conducted frequently (generally, at least once a day) with a full backup taking place on a routine basis.

Backups can be stored on many different media. These media types may include the following:

- Diskettes;

- Tape Backup Devices;

- Removable Disk Cartridges; and

- Rewritable CD.

The current backup and recovery requirements, and acceptable downtimes for Title IV systems range from daily, weekly, quarterly, annually, or on request depending on the system. The most predominant backup media is tape that is retained for one week, one month, to one year, depending on the system.

Based upon procedures for current Title IV systems, as well as industry standards and best practices, the following requirements for backup procedures are recommended:

- Files shall be backed up incrementally every day;
- Files shall be backed up fully at least once a week;
- Files may require full backup periodically based upon processing cycles, e.g. after month-end processing;
- Daily backups shall be retained for at least one week;
- Weekly backups shall be retained for at least one month; and
- Other periodic (e.g. month-end) backups shall be retained for one year.

4.4.6 Recovery Requirements

Faulkner Information Services indicates that 50 percent of businesses that lose their data due to disasters go out of business within 24 months and, according to the US Bureau of Labor, 93 percent are out of business within five years. No business should be without a disaster recovery plan.

The EASI/ED system will require a disaster recovery plan to bring the system back to full operating capacity in the event of a disaster, or some event that causes the system to fail. This disaster recovery plan should include the following:

- **Data center backup and user recovery procedures** - The disaster recovery plan should take into account the protection of the central data center functions and the departmental functions distributed throughout the organization.
  - A hot site is a complete data center equipped with processors, peripherals, and communications equipment that an organization can use on a contractual basis. A hot site service offers almost uninterrupted availability of a functional data center. Within a few hours, the essential operations can be initiated, and business can resume.
  - A cold site is an empty room or building with a raised floor, air-conditioning units, and power supplies. It contains no computers or peripherals, although it may be equipped with dial-up lines for a communications network. A cold site is best utilized in conjunction with a hot site. The hot site is used while new equipment is shipped and installed at the cold site.
- **Evacuation and emergency (first aid) procedures** - A recovery plan should provide specific procedures to anticipate problems ranging from data loss or partial file destruction to complete facility destruction. The procedures should include names, addresses, and telephone numbers of all personnel who must be present in a recovery situation:
  - Public relations, financial, legal, and administrative recoveryprocedures and personnel;[-
  - List of backup personnel in case the primary resources are not available;

- Contact information for restoring the building and rewiring;
- Identification of critical products and services;
- Determination of how long the enterprise can afford to be down. There should be a mirror site where all files are being shadowed, or some other offsite storage; and
- Analysis of risks and exposures.

A calendar recovery plan to prioritize applications - The calendar lists all time-sensitive procedures, such as month-end reports, which may take higher priority in the event of a disaster. Standard recovery procedures also should be implemented in specific order. Vital information about each application should be recorded and include the following:

- Potential cost of not running each application;
- Names and locations of various data files and jobs associated with the application;
- Run-time information such as system overhead and workload;
- System configuration;
- System and application documentation and system support material (e.g. pre-printed forms). This material should be stored in a secure location off-site;
- Pertinent user information;
- Estimate of the minimum system requirements for the most critical applications; also an overall minimum configuration. As the system comes back up after a disaster, the non-critical applications will eventually need to be run as well. This will likely require more power than is available from the recovery configuration, so additional system resources [?] may be needed;
- Escalation procedures for the more serious business interruption scenarios. Management should immediately notify vendors of problems and alert recovery teams if a timely resolution does not take place;
- Detailed tactical recovery documents. The documents should be copied and kept in several offsite locations, preferably in the hands of those who must implement it; and
- Determination of which primary vendors and other contacts to alert first.

- **Data recovery requirements: -** The maximum time between failure and recovery of data shall be one day**.**

4.4.7 Availability Requirements

System availability measures a system's ability to provide access. A system that is highly available provides a user with immediate access to services, whereas a system that suffers from

low availability requires users to wait. An analogy of this idea is when a user picks up a telephone handset to place a telephone call. If there is a dial tone, the telephone line is available and the user may place a call. However, if there is no dial tone, the telephone line is not available and the user must wait. If the telephone line does not become available, the user will attempt to isolate the problem or find another telephone.

Availability measures should identify when the service would be open to users. A common misconception is that availability is always related to performance. These two concepts are not interdependent. A busy network may be unusable because of slowness, but all resources remain available.

An organization must be able to account for both planned and unplanned system unavailability and define the acceptable levels of performance. This is a critical factor in determining the success of an information system and many organizations are now ensuring that availability measures are included in any contractual agreement they undertake with a service provider.

Unfortunately, there are no real industry standards for information systems (e.g., local area networks, wide area networks, Internet, Intranet, etc.) availability at this time. For example, wide area networks, by their very nature, are sporadic which makes them difficult to monitor and measure. However, the de facto standard for availability rates for most information systems generally range from 95 – 100 percent, with the concentration being at the 98 - 99.9 percent levels.

Where information was available, the required system operation times and required availability for the current Title IV Systems are listed in Table 4 - 4 below. Availability estimates identified in the current Title IV system documentation (e.g., performance work statments, user documentation, Project EASI/ED documentation) provide a solid baseline for the minimum levels of the EASI/ED system. ED's current requirements for system availability range from 98 – 99.98 percent, with a maximum of 30 hours of allowable maintenance (downtime) per month.

Current Title IV System	Required Hours of Operation	Required Availability
CBS	8am - 8pm, M-F	N/A
CDS	8am - 8pm, M-F	N/A
CPS	8am - 8pm, M-F	N/A
FFELP	8am - 10pm, M-Sat	N/A
LOS	24 hours, 7 days	99.75%
LSS	7am - 10pm, M-Sat	99%
NSLDS	24 hours, 7 days	99.75%
RFMS	N/A	98%
TIVWAN	24 hours, 7 days	99.98%

Table 4 -4: Current Title IV Systems Operation Times and Required Availability

Based on de facto standards and best practices, with an eye towards current Title IV practices, the following specification represent the availability requirements for Project EASI/ED.

- The system shall maintain a minimum level of availability of 99.98 percent.

4.4.8 Reliability Requirements

Reliability addresses the issue of a system's ability to continue operating despite the failure of some critical component or element, and the creditability of the information provided. Typically,

reliability involves implementing redundancy which can increase component and media costs. The terms reliability and availability are often used synonymously, but for the purposes of this document, reliability is defined as it is in the TIVWAN PWS.

Of the current Title IV systems in use, only the TIVWAN possesses any specific requirements for reliability. Its PWS indicates that reliability shall be measured through two performance criteria: Access Denial Probability and Bit Error Rate. Access Denial Probability is defined as the ratio of total access attempts that result in access denial, to total system access attempts. The performance work statement calls for an Access Denial Probability of less than 1 percent (or, 99 percent reliability). Bit Error Rate is defined as the ratio of total incorrect, lost, or extra user data bits (i.e., user submitted information) received at the destination to total user data bits submitted at the source. In other words, it is a determination of how much of what is sent is received correctly at its proper destination. The performance work statement calls for only a one in one trillion probability of misrouting information to the wrong end user.

In order to increase reliability, the amount of time required to fix any failures or problems within the system must be minimized. The repair interval, Mean Time to Repair (MTTR), is the average length of time required to correct a fault when the system or a component of the system fails. The de facto industry standard for MTTR is approximately four hours. The MTTR of newer systems is expected to be much lower than that of previous generations of systems, due to the widespread use of fault-isolation capabilities that prevent further system deterioration.

While four hours is the industry standard for MTTR, other quantitative numbers for reliability are not as easily identified. Similar to availability, there are no real industry standards for information system reliability at this time. However, de facto standards for reliability levels generally range from 98 – 100 percent, with the concentration being 98.5 - 100 percent.

Based on de facto standards and maintaining at least current system performance, the following specifications represent the reliability requirements for Project EASI/ED.

- The system shall maintain reliability of systems, as measured through Access Denial Probability, of at least 99 percent.

- The system, whether it be a WAN or Virtual Private Network (VPN), shall have a one in one trillion probability of misrouting information to the wrong end user.

- The Mean Time To Repair of the system must not exceed four hours.

4.5 High-Level Performance Issues for Reused Title IV Systems

This subsection will briefly discuss issues regarding current Title IV systems and their ability to comply with the performance requirements developed in subsection 4.4. It examines systems that can be confirmed as non-reuse candidates, and looks at transaction and volume information that proposed reuse candidates must accommodate.

The LSS system and the FFEL system can be confirmed as two systems that are not viable reuse candidates. PSS staff have indicated that LSS and the Debt Collection Subsystem (DCS) of FFEL will not be able to handle the volumes of transactions that are expected by the year 2003, if not sooner. Given the inability of these systems to handle the expected volumes, neither of these two systems can be considered reuse candidates.

Four Title IV systems are documented in the *Project EASI/ED Transition Strategy* as candidates for substantial reuse as part of the EASI/ED implementation. These systems are CPS, LCS, NSLDS, and RFMS.

However, in order to be viable reuse candidates, they must be able to accommodate expected transaction volumes. For example, the CPS system must handle the expected volume of ten million FAFSAs in the year 2001, including possibly two million of these via the Web.

Similarly, it is anticipated that when RFMS becomes fully operational, between 23 and 30 million Pell Grant disbursement records may need to be processed annually. Use of RFMS to process Direct Loan originations and disbursements in addition to Pell must clearly be contingent on its ability to successfully meet these increased performance demands.

Appendix F presents estimated transaction volumes for a number of performance-sensitive transactions that these systems must support.